# Stochastic Image Denoising Based on Markov-Chain Monte Carlo Sampling

Alexander Wong, Akshaya Mishra, Wen Zhang, Paul Fieguth, and David A. Clausi

*Vision and Image Processing (VIP) research group*
*Department of Systems Design Engineering, University of Waterloo*
*Waterloo, Canada N2L 3G1*
*{a28wong,akmishra,wxzhang,pfieguth,dclausi}@uwaterloo.ca*

**Abstract**

A novel stochastic approach based on Markov-Chain Monte Carlo sampling is investigated for the purpose of image denoising. The additive image denoising problem is formulated as a Bayesian least squares problem, where the goal is to estimate the denoised image given the noisy image as the measurement and an estimated posterior. The posterior is estimated using a nonparametric importance-weighted Markov-Chain Monte Carlo sampling approach based on an adaptive Geman-McClure objective function. By learning the posterior in a nonparametric manner, the proposed Markov-Chain Monte Carlo denoising (MCMCD) approach adapts in a flexible manner to the underlying image and noise statistics. Furthermore, the computational complexity of MCMCD is relatively low when compared to other published methods with similar denoising performance. The effectiveness of the MCMCD method at image denoising was investigated using additive Gaussian noise, and was found to achieve state-of-the-art denoising performance in terms of both peak signal-to-noise ratio (PSNR) and mean structural similarity (SSIM) metrics

when compared to other published methods.

*Keywords:* image denoising, Markov-Chain Monte Carlo.

---

## 1. Introduction

One of the fundamental challenges in the field of image processing and computer vision is image denoising, where the underlying goal is to produce an estimate of the original image by suppressing noise from a noise-contaminated version of the image. Image noise may be caused by different intrinsic (i.e., sensor) and extrinsic (i.e., environment) conditions which are often not possible to avoid. Therefore, image denoising plays an important role in a wide range of applications such as photo restoration, visual tracking, and image segmentation, where obtaining the original image content is crucial. While many algorithms have been proposed for the purpose of image denoising, the problem of image noise suppression remains an open challenge, especially in situations where the images are acquired under poor conditions.

Image denoising algorithms can generally be categorized as either transform domain methods or spatial domain methods. In transform domain methods, noise suppression is performed based on the characteristics of the coefficients in a different domain (e.g., the Fourier, wavelet, or discrete cosine). Transform domain methods can be further divided into global methods, such as Fourier Wiener filtering [1], and local methods, such as empirical Wiener filtering [2], Gaussian scale mixture denoising [3], wavelet shrinkage [4, 5, 6, 7, 8, 9], and shape-adaptive discrete cosine transform filtering [10].

Spatial domain methods take advantage of the spatial information re-

dundancy inherent in images to suppress image noise and can be further categorized as either local or global methods. A majority of existing denoising methods take advantage of information redundancy within small local neighborhoods to suppress image noise. These methods include traditional denoising methods such as box filtering and Gaussian filtering [11], and detail-preserving adaptive denoising methods such as Lee filtering [12], total variation [13], Bayesian estimation [14, 15], anisotropic filtering [16], bilateral filtering [17, 18, 19], and trilateral filtering [20]. The main advantage of local spatial domain methods is that they are relatively computationally efficient compared to global spatial domain methods. However, such methods generally perform poorly on images contaminated by a high level of noise since the local information redundancy is insufficient to provide a good noise-free image estimate.

To address the shortcomings of local methods, non-local spatial domain approaches have been introduced that exploit information redundancy over the whole image [21, 22, 23, 24]. The main advantage of such non-local methods is that they are more effective at handling situations characterized by low signal-to-noise ratio due to the high information redundancy found within the entire image. However, despite performance optimizations that have been introduced [22, 23], such methods remain computationally expensive due to the amount of information that must be processed in obtaining the image estimate.

In this study, we investigate an alternative approach to the problem of image denoising based on stochastic optimization via Markov-Chain Monte Carlo sampling. By formulating the problem as a Bayesian least squares

problem and taking a nonparametric stochastic approach to solving this problem, the proposed Markov-Chain Monte Carlo denoising (MCMCD) approach adapts to the underlying image and noise statistics in a flexible manner. The presented MCMCD approach has relatively low computational complexity when compared to published methods with similar denoising performance.

Given the wealth of literature on the use of Markov-Chain Monte Carlo (MCMC) sampling for the purpose of image restoration, it is important to distinguish the proposed MCMCD approach from other image denoising methods that utilize stochastic sampling concepts. In the seminal work by Geman and Geman [14], the problem of image restoration is formulated as a maximum a posteriori (MAP) problem, where the goal is to maximize the conditional probability of the original image given the degraded image. An iterative simulated annealing approach is then employed to solve this problem based on a Markov Random Field (MRF), where random local changes are made to the image according to a local conditional probability distribution, constructed using a Markov-Chain Monte Carlo sampling approach.

The use of MCMC sampling in the classical approach is fundamentally different from the proposed method for two important reasons. First, the classical approach employs MCMC sampling to randomly construct possible configurations of the restored image from a posterior distribution defining all possible spatial-intensity configurations indirectly. In contrast, the proposed method employs MCMC sampling to construct an estimate of the posterior distribution defining all possible intensity configurations on a per-pixel basis. As such, the classical approach determines the restored image iteratively as

the configuration that maximized the posterior, while the proposed method computes the restored image in a non-iterative manner based on the constructed posterior distribution for each pixel, a distinction which leads to the lower computational complexity of the proposed method compared to the classical approach. Second, while the classical approach encodes the local spatial context in a local MRF framework for the MCMC sampling process, the proposed method takes into account local spatial context within a global framework, allowing the proposed method to take greater advantage of information throughout the image than the classical approach.

In wavelet denoising approaches that employ MCMC sampling [8, 9], an MRF of local wavelet coefficients is constructed in a stochastic manner. MCMC sampling in the wavelet approach is fundamentally different from the proposed method because the posterior distribution being estimated by the wavelet approach relates to a binary decision about whether a wavelet coefficient is noise-free or not, while the posterior distribution being estimated by the proposed method relates to the underlying intensity at a particular pixel in the image.

The paper is organized as follows. The theory behind MCMCD is presented in Section 2. A summary of the MCMCD algorithm is presented in Section 3. The computational complexity of MCMCD is examined in Section 4. The testing methods and experimental results are discussed in in Section 5. Finally, conclusions are drawn and future work is discussed in Section 6.

## 2. Theory

The additive image noise problem formulation and the corresponding Bayesian least squares solution formulation is presented in Section 2.1. The Markov-Chain Monte Carlo sampling scheme is described in Section 2.2.

### *2.1. Problem formulation*

Let $S$ be the set of sites in a discrete lattice $\pounds$ upon which an image is defined and $s \in S$ be a site in $\pounds$. Let the measured noise-contaminated image $F = \{F(s)|s \in S\}$, noise-free image $G = \{G(s)|s \in S\}$, and noise $N = \{N(s)|s \in S\}$ be random fields on $S$. The relationship between $F$, $G$, and $N$ in an additive image noise model can be generally expressed as

$$F(s) = G(s) + N(s), \tag{1}$$

Given Eq. (1), the computation of $G(s)$ can be viewed as an inverse problem and hence be solved as an estimation problem. Formulating the inverse problem as a Bayesian least squares estimation problem [25], given the measurement $F(s)$, the state estimate of $G(s)$ can be expressed as

$$\begin{aligned} \hat{G}(s) = & \underset{\hat{G}(s)}{\arg\min} \left\{ E\left( \left( G(s) - \hat{G}(s) \right)^2 | F(s) \right) \right\} \\ = & \underset{\hat{G}(s)}{\arg\min} \left\{ \int \left( G(s) - \hat{G}(s) \right)^2 p\left( G(s) | F(s) \right) dG(s) \right\}, \end{aligned} \tag{2}$$

where $p(G(s)|F(s))$ is the posterior. Given Eq. (2), the analytical solution can be derived as follows. Taking the derivative of Eq. (2) gives

$$\begin{aligned} & \tfrac{\partial}{\partial \hat{G}(s)} \int \left[ G(s) - \hat{G}(s) \right]^2 p(G(s)|F(s)) dG(s) \\ = & \int \left\{ -2 \left[ G(s) - \hat{G}(s) \right] p(G(s)|F(s)) dG(s) \right\}. \end{aligned} \tag{3}$$

Setting the derivative to zero, an expansion of Eq. (3) can be expressed as

$$\int \hat{G}(s)p(G(s)|F(s))dG(s) = \int G(s)p(G(s)|F(s))dG(s). \tag{4}$$

Given that

$$\begin{aligned} \int \hat{G}(s)p(G(s)|F(s))dG(s) &= \hat{G}(s) \int p(G(s)|F(s))dG(s) \\ &= \hat{G}(s), \end{aligned} \tag{5}$$

Eq. (4) simplifies into

$$\hat{G}\left(s\right) = \underbrace{\int G\left(s\right) p\left(G\left(s\right)|F\left(s\right)\right)dG\left(s\right)}_{E(G(s)|F(s))}, \tag{6}$$

which implies that the optimal estimate of $G\left(s\right)$ is $E\left(G\left(s\right)|F\left(s\right)\right)$, the conditional mean of $G\left(s\right)$ on $F\left(s\right)$. Unfortunately, the conditional mean of $\hat{G}\left(s\right)$ can be a highly complicated and nonlinear function of $F\left(s\right)$ and difficult to solve in an analytical manner. To work around this issue, typical strategies for computing Eq. (6) include approximating the problem as a simpler Bayesian linear least squares problem [26], as well as modeling the posterior $p(G(s)|F(s))$ using parametric statistical models [27]. However, such strategies often result in poor state estimates for the image denoising problem given the complex, nonlinear nature of images.

To improve the quality of the state estimate of $G\left(s\right)$, a different strategy is explored in this paper to tackle the problem of modeling the posterior. Given the measurement $F(s)$, we investigate the potential for computing a nonparametric approximation of $p\left(G(s)|F(s)\right)$, denoted as $\hat{p}\left(G(s)|F(s)\right)$, using an importance-weighted Markov-Chain Monte Carlo posterior estimation strategy that takes into account local spatial-intensity context within a global

framework. The use of local spatial-intensity context is important as it allows for improved discrimination between the underlying noise-free image $G(s)$, which is typically characterized by strong local spatial-intensity interactions, and the noise $N(s)$, which typically exhibits very low local spatial-intensity interactions.

### *2.2. Markov-Chain Monte Carlo posterior estimation*

The importance-weighted Markov-Chain Monte Carlo posterior estimation strategy for computing $\hat{p}\left(G(s)|F(s)\right)$ can be described as follows. Let $s$ be the site for which the posterior is being estimated. In Markov-Chain Monte Carlo density estimation [28], an unknown target distribution (in this case, the posterior) is estimated indirectly by drawing a sequence of samples $\Omega = \{s_0, s_1, \ldots, s_\eta\}$ from a known distribution $Q(s')$ that are likely to be realizations of $p(G(s)|F(s))$. To initialize the sequence of samples $\Omega$, let $s_0 = s$. At the $k^{\text{th}}$ iteration, a candidate sample $s'_k$ is drawn from a known instrumental distribution $Q(s'_k|s_{k-1})$. There are two main general criteria in the selection of a suitable instrumental distribution for the purpose of denoising. First, the chosen distribution should promote the sampling of sites within close proximity to $s_k$, as such sites are more likely to share similar local spatial-intensity characteristics as $s_k$ and hence provides meaningful information for estimating $G\left(s\right)$. Second, the chosen distribution should not completely prevent other sites within the image from being sampled, as they may also contain important information. Based on these two criteria, a convenient choice for the instrumental distribution $Q(s'_k|s_{k-1})$ is a Gaussian distribution centered at $s_{k-1}$, since such a distribution would allow the probability of sampling to monotonically decrease as we move away from $s_{k-1}$,

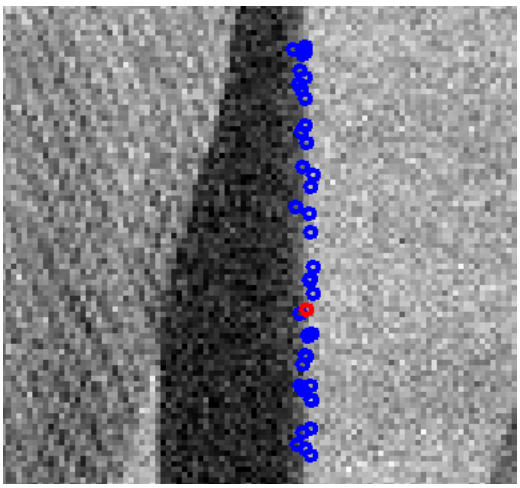


Figure 1: An example of the samples accepted in $\Omega$ (indicated by blue circles) for a particular site $s$ (indicated by the red circle). An important observation that can be made from the locations of the samples is that the Markov-Chain Monte Carlo sampling "walks" along sites with similar image characteristics that are relevant to the posterior estimation. Hence, this approach to sampling takes advantage of the global framework without the computational burden of evaluating all sites on $S$.

thus encouraging increased sampling of sites within close proximity while still allowing sites farther away to still be sampled to better utilize global information within the image. Based on testing, it was found that the use of a Gaussian instrumental distribution provided high quality estimates of $G(s)$, although other instrumental distributions may also be used. For MCMCD, the Gaussian instrumental distribution $Q(s'_k|s_{k-1})$ can be defined as

$$Q(s'_k|s_{k-1}) = \frac{1}{2\pi\sigma_s} \exp\left[-\left(\frac{(s'_k - s_{k-1})^2}{2\sigma_s^2}\right)\right], \tag{7}$$

where $\sigma_s$ represents the spatial variance of $Q(s'_k|s_{k-1})$.

To determine whether the sample candidate $s'_k$ is accepted as part of $\Omega$,

the acceptance probability of $s'_k$ given $s_{k-1}$ is given by

$$\alpha\left(s'_k|s_{k-1}\right) = \min\left\{1, \frac{\phi(s'_k|s_0)}{\phi(s_{k-1}|s_0)}\right\} \tag{8}$$

where $\phi(s'_k|s_0)$ is an objective function that approximates the likelihood that a sample is a realization of $p(G(s)|F(s))$. To take into account the local spatial-intensity context inherent in the image, as well as the underlying noise and image statistics, we propose a novel adaptive objective function that evaluates the established and robust Geman-McClure error statistics [29] between the local neighborhoods around the samples $\rho_{s_0}$ and $\rho_{s'_k}$,

$$\phi\left(s'_k|s_0\right) = \prod_j exp\left[-\frac{(F^{\rho_{s'_k}}(j) - F^{\rho_{s_0}}(j))^2}{\frac{\sigma_n^4}{\sigma_l^2} + (F^{\rho_{s'_k}}(j) - F^{\rho_{s_0}}(j))^2}\right]. \tag{9}$$

where $F^{\rho_{s'_k}}(j)$ and $F^{\rho_{s_0}}(j)$ are the intensities of the $j^{\text{th}}$ site in the local neighborhoods around $s'_k$ and $s_0$, respectively, and $\sigma_n$ and $\sigma_l$ are the noise variance and local variance of $\rho_{s_0}$, respectively. The objective function in Eq. (9) gives greater values to sites with similar local spatial-intensity relationships as $s_0$, as they are more likely to belong to the same distribution as $s$. The sample candidate $s'_k$ is then accepted into $\Omega$ with a probability of $\alpha\left(s'_k|s_0\right)$. This probabilistic acceptance step is realized in the proposed algorithm as follows. First, a value $u$ is randomly sampled from a uniform distribution $U(0,1)$. If $u > \alpha\left(s'_k|s_0\right)$, then the sample candidate $s'_k$ is discarded. However, if $u \leq \alpha\left(s'_k|s_0\right)$, then $s'_k$ is included into the sequence of samples $\Omega$.

The sampling process is repeated until the desired number of iterations $m$ have been met. The selection of $m$ becomes a tradeoff between computational performance (e.g., fewer iterations) and image quality (e.g., more iterations). However, based on testing, it was found that setting $m$ to values greater than

200 yields little additional benefit in terms of image quality, making $m = 200$ a good compromise between computational performance and image quality. An example of the samples accepted in $\Omega$ for a particular site is shown in Fig. 1. An important observation that can be made from the locations of the samples is that the Markov-Chain Monte Carlo sampling "walks" along sites with similar image characteristics that are relevant to the posterior estimation. Hence, this approach to sampling takes advantage of the global framework without the computational burden of evaluating all sites on $S$.

Furthermore, due to the stochastic acceptance criterion, the number of accepted samples $\eta$ can vary from pixel to pixel. An example of the variability in $\eta$ for the $512 \times 512$ "Lena" image contaminated by Gaussian noise with $\sigma_n = 30$, with a section of the image and corresponding sampling density map is shown in Fig. 2. The sample density is higher in areas of high local spatial-intensity similarities (e.g., homogeneous regions) and lower in areas with more unique local spatial-intensity characteristics (e.g., structured regions).

Given $\Omega$, the posterior estimate $\hat{p}\left(G(s)|F(s)\right)$ can be computed based on the concept of importance-weighted Markov-Chain Monte Carlo posterior estimation [30]. Let the set of objective values computed in Eq. (9) be used as the associated importance weights defined on the set of samples given $s$,

$$\Phi = \left\{\phi\left(s'_0|s_0\right), \phi\left(s'_1|s_0\right), \ldots, \phi\left(s'_\eta|s_0\right)\right\}. \tag{10}$$

## 3. Summary of proposed method

To provide a clearer understanding of how the theory presented in Section 2 can be applied in practice, MCMCD can be summarized as follows:

1. At each site $s$, randomly draw a sample $s'_k$ from instrumental distribution $Q(s'_k|s_{k-1})$ in Eq. (7).
2. Compute acceptance probability $\alpha\left(s'_k|s_{k-1}\right)$ in Eq. (8).
3. Generate a random value $u$ according to a uniform distribution $U(0,1)$.
4. If $u \leq \alpha\left(s'_k|s_{k-1}\right)$, include $s'_k$ into the sequence of samples $\Omega$. Otherwise, discard $s'_k$.
5. Repeat steps 1 to 4 until the desired number of iterations $m$ is met.
6. Estimate posterior $\hat{p}\left(G(s)|F(s)\right)$ based on the set of drawn samples $\Omega$ in Eq. (**??**).
7. Compute the noise-free image estimate $\hat{G}(s)$ as the discrete form of the conditional mean in Eq. (6).

## 4. Computational complexity

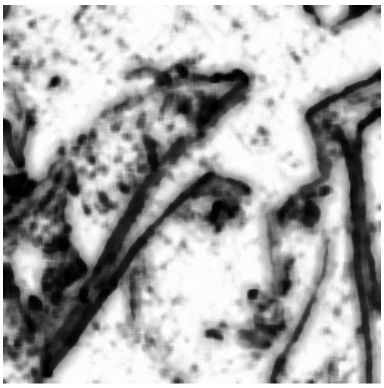
The general time complexity of MCMCD is $O(n)$ thus the run-time of the algorithm scales linearly with the size of the image. However, given the stochastic nature of the Markov-Chain Monte Carlo posterior estimation process, which is the most computational complex aspect of MCMCD, the computational complexity of the algorithm can vary depending on the underlying image and noise characteristics. Nevertheless, the theoretical per-pixel computational complexity upper-bound of MCMCD in an unoptimized form can be expressed as

$$n_\rho \times m, \tag{11}$$

where $n_\rho$ is the number of pixels in a local region. In this case, the number of iterations $m$ is 200 and $n_\rho = 49$, making the per-pixel computational

**Noisy ($F(s)$)**

**Denoised ($\hat{G}(s)$)**



**Sample density ($\eta$)**

Figure 2: "Lena" image contaminated by additive Gaussian noise with $\sigma_n = 30$ and the corresponding sample density map and denoised image. Higher intensity values in the sample density map indicate that more samples were used to estimate the denoised image at that particular pixel. The sample density is higher in areas of high local spatial-intensity similarities (e.g., homogeneous regions) and lower in areas with more unique local spatial-intensity characteristics (e.g., structured areas).

complexity of MCMCD for this particular case approximately

$$n_\rho \times m = 49 \times 200 = 9800. \tag{12}$$

To put this into perspective, the per-pixel computational complexity of non-local means [22] is

$$n_\rho \times n_s, \tag{13}$$

where $n_s$ is the number of pixels in the search window. Using a search window size of $21 \times 21$ and $7 \times 7$ regions [22], the per-pixel computational complexity of non-local means is approximately

$$n_\rho \times n_s = 49 \times 441 = 21609. \tag{14}$$

Furthermore, based on the computational complexity analysis presented in [24], the per-pixel computational complexity of the state-of-the-art BM3D [24] algorithm in an unoptimized form is

$$59488 + 22C_{t2d} + ((128/3)C_{t1d}), \tag{15}$$

where $C_{t1d}$ and $C_{t2d}$ are the computational complexity of the 1D and 2D transforms (e.g., DCT, DST, or wavelet transform). As such, discounting algorithmic optimizations, MCMCD has a relatively low per-pixel computational complexity (9800) when compared to these state-of-the-art methods (21609 for non-local means, and greater than 59488 for BM3D). From a practical perspective, the MCMCD method was implemented and tested on an Intel Pentium 4 3 GHz machine with 1 GB of RAM, taking approximately 20 seconds to denoise a 512×512 image. The speed of MCMCD can be improved significantly through optimized code and parallel processing paradigms. Finally, from an implementation complexity perspective, the

concept of MCMCD is relatively straightforward to implement based on the algorithm presented in Section 3.

## 5. Experimental results

In this section, two sets of experiments are performed to illustrate the performance of the proposed MCMCD method. In the first set of experiments, MCMCD is compared with the classical MCMC image restoration approach [14] to demonstrate the advantage of proposed MCMCD approach over the classical approach in terms of the visual quality of the denoised images produced. In the second set of experiments, MCMCD is compared to existing state-of-the-art denoising methods under different noise levels to assess the overall performance of MCMCD both quantitatively and qualitatively. Standard test images such as Peppers, Lena, Barbara, and Hill were used, as shown in Fig. 3, which is consistent with existing literature [24]. All images are 512×512 in resolution with the exception of "Peppers", which is 256×256 as presented in [24].

### *5.1. Comparison with classical MCMC approach*

To evaluate the performance gains of the proposed MCMCD method when compared with the classical MCMC image restoration approach [14], qualitative comparative analysis was performed for the standard Barbara and Hill test images. For each test image, Gaussian noise was added with $\sigma_n = 20$. Figs. 4 and 5 show the "Hill" and "Barbara" images contaminated by additive Gaussian noise with $\sigma_n = 20$ and the corresponding denoised images using the classical approach and the proposed MCMCD method. While the classical approach is able to well preserve image detail for the most part,

Figure 3: Set of test images (from top to bottom, left to right): Peppers, Barbara, Lena, and Hill.

there is noticeable degradation in finer scale details such as the windows in the "Hill" image and the stripes in the "Barbara" image. Furthermore, while much of the noise has been suppressed, the denoised images produced by the classical approach exhibits noticeable artifacts in both cases. On the other hand, MCMCD is able to effectively suppress noise in these images without introducing artifacts while preserving image detail, which is particularly noticeable in the windows in the "Hill" image and in the table cloth and stripes in the "Barbara" image.

*5.2. Comparison with state-of-the-art methods*

To evaluate the potential of the proposed MCMCD method for image denoising when compared with several state-of-the-art denoising methods, qualitative and quantitative comparative analyses were performed for different noise levels. For each test image, Gaussian noise was added with $\sigma_n = 20, 30, 40, 50$. The other denoising algorithms tested were the state-of-the-art collaborative non-local filtering **(BM3D)** [24], NeighShrink SURE wavelet denoising **(NS)** [7], and non-local means denoising **(NLM)** [22]. These methods were chosen as they represent the recent literature in transform-based methods (NS), and non-local spatial domain methods (NLM and BM3D). The tested methods are configured based on parameters presented in their respective literatures.

For quantitative comparison purposes, the peak signal-to-noise ratio (PSNR) was computed according to the formula,

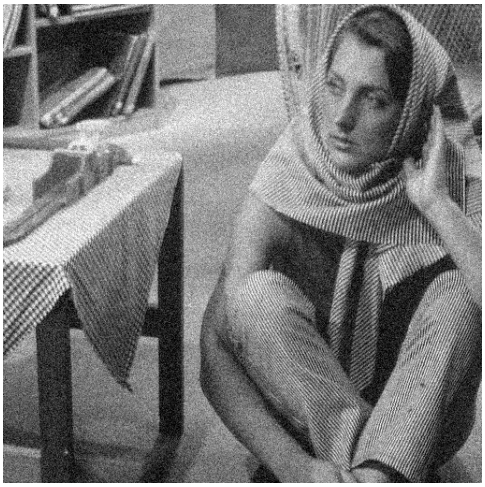
$$PSNR\left(\hat{G}\right) = 10\log_{10}\left(\frac{G_{MAX}^2}{\frac{1}{N}\sum\left(\hat{G}(s) - G(s)\right)^2}\right), \tag{16}$$

**Noisy**

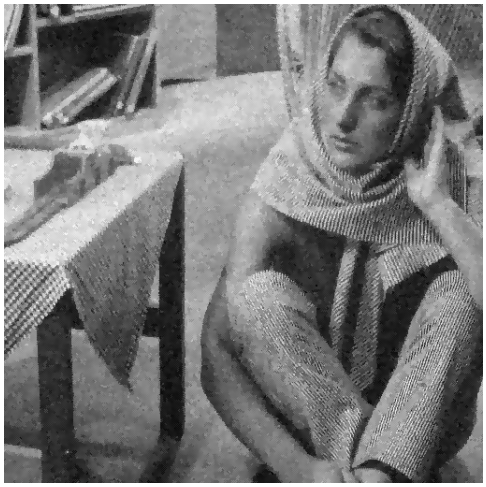
**Classical approach**

**MCMCD**

Figure 4: "Hill" image contaminated by additive Gaussian noise with $\sigma_n = 20$ and the corresponding denoised image using the classical MCMC approach (top right) and MCMCD (bottom).



**Noisy**



**Classical approach**

**MCMCD**

Figure 5: "Barbara" image contaminated by additive Gaussian noise with $\sigma_n = 20$ and the corresponding denoised image using the classical MCMC approach (top right) and MCMCD (bottom).

where $G_{MAX}$ is the maximum possible value of $G$, and $N$ is the total number of pixels in the image. In addition to the peak signal-to-noise ratio, image quality was also assessed using mean structural similarity (MSSIM) [32], which has been widely used as an alternative metric to PSNR as an image quality assessment metric. For MCMCD, $\sigma_s = 21$ (Eq. (7)), circular regions with radius of 3 (Eq. (9)), and $m = 200$ were used for all tests as they provide strong denoising performance based on testing.

Table I shows the PSNR and SSIM values at different noise levels for the tested methods, and two key insights can be observed. First, MCMCD is competitive with the tested state-of-the-art methods, producing better PSNR and SSIM values than NS and NLM across all noise levels, albeit lower PSNR and SSIM values when compared to BM3D. Second, the noise suppression performance of MCMCD decreases as $\sigma_n$ increases. This is due to the fact that the spatial-intensity relationships of very fine image details in the test images become almost indistinguishable from that of the noise for situations characterized by low signal-to-noise ratios. Under these situations, identifying samples that are likely to be realizations of the noise-free image distribution is difficult. Therefore, more irrelevant samples are used under these situations which reduces the quality of the denoised image.

Fig. 6 shows the "Barbara" image contaminated by additive Gaussian noise with $\sigma_n = 30$ and the corresponding denoised images using the tested methods. Based on visual inspection, all tested methods are effective at suppressing image noise, with a zoomed-in region extracted from each of the denoised images in Fig. 7. While NS provides good image detail preservation, significant artifacts are visible in the produced image and as such degrade

Table 1: PSNR and MSSIM results with respect to noise level ($\sigma$). MSSIM results are shown in brackets.

| **Method** | $\sigma$**=20** | $\sigma$**=30** | $\sigma$**=40** | $\sigma$**=50** |
|---|---|---|---|---|
| **Peppers** | | | | |
| Noisy | 22.22 (0.42) | 18.56 (0.29) | 16.09 (0.21) | 14.13 (0.16) |
| BM3D [24] | 31.21 (0.88) | 29.27 (0.85) | 27.71 (0.81) | 26.35 (0.77) |
| NS [7] | 29.14 (0.81) | 26.90 (0.76) | 25.23 (0.69) | 24.04 (0.67) |
| NLM [22] | 28.77 (0.83) | 26.53 (0.78) | 24.85 (0.74) | 23.43 (0.70) |
| MCMCD | 30.44 (0.86) | 28.40 (0.81) | 26.72 (0.78) | 25.34 (0.72) |
| **Barbara** | | | | |
| Noisy | 22.10 (0.47) | 18.59 (0.34) | 16.09 (0.25) | 14.14 (0.19) |
| BM3D [24] | 31.76 (0.90) | 29.75 (0.86) | 27.97 (0.82) | 27.10 (0.78) |
| NS [7] | 28.94 (0.84) | 26.79 (0.77) | 25.46 (0.72) | 24.41 (0.68) |
| NLM [22] | 29.23 (0.84) | 26.75 (0.77) | 25.10 (0.70) | 23.96 (0.65) |
| MCMCD | 30.53 (0.89) | 28.76 (0.82) | 26.48 (0.75) | 25.82 (0.71) |
| **Lena** | | | | |
| Noisy | 22.11 (0.34) | 18.58 (0.21) | 16.11 (0.15) | 14.13 (0.16) |
| BM3D [24] | 33.04 (0.87) | 31.24 (0.84) | 29.77 (0.81) | 28.77 (0.78) |
| NS [7] | 31.09 (0.84) | 29.14 (0.80) | 27.91 (0.77) | 27.04 (0.73) |
| NLM [22] | 29.13 (0.79) | 27.30 (0.75) | 26.18 (0.72) | 26.51 (0.72) |
| MCMCD | 32.18 (0.86) | 30.05 (0.82) | 28.86 (0.76) | 27.61 (0.75) |
| **Hill** | | | | |
| Noisy | 22.09 (0.41) | 18.56 (0.26) | 16.10 (0.17) | 14.14 (0.13) |
| BM3D [24] | 30.51 (0.79) | 28.98 (0.74) | 27.74 (0.68) | 27.03 (0.66) |
| NS [7] | 29.05 (0.74) | 27.35 (0.68) | 26.27 (0.63) | 25.67 (0.60) |
| NLM [22] | 26.57 (0.62) | 25.45 (0.57) | 24.86 (0.54) | 24.61 (0.53) |
| MCMCD | 29.83 (0.77) | 28.02 (0.70) | 26.86 (0.63) | 25.94 (0.62) |

the overall visual quality of the image. The image produced by NLM exhibits significantly fewer artifacts than NS, but does not preserve fine details such as the stripe patterns in Barbara's clothes as well as BM3D and MCMCD. The images produced by BM3D and MCMCD both exhibit better image detail preservation than NLM while introducing significant fewer artifacts into the image when compared to NS.

Finally, particularly noticeable in the zoomed-in regions shown in Fig. 7, the results produced by MCMCD are visually comparable to BM3D, with BM3D providing slightly better detail preservation while MCMCD exhibits slightly weaker artifacts in the homogeneous regions. Fig. 8 shows the "Lena" image contaminated by additive Gaussian noise with $\sigma_n = 40$ and the corresponding denoised images using the tested methods. Based on visual inspection, the observations made from the "Barbara" image also hold true for the "Lena" image, where MCMCD and BM3D provides the best image quality, with BM3D providing better detail preservation while MCMCD exhibits weaker artifacts. It is important to note that MCMCD is less computational complex than BM3D, as discussed in Section 4.

## 6. Conclusions

In this study, the use of stochastic optimization via Markov-Chain Monte Carlo sampling is explored for the purpose of image denoising. The proposed MCMCD approach allows the additive denoising problem to be solved in a flexible, nonparametric fashion by adapting to the underlying image and noise statistics. Furthermore, the proposed MCMCD approach has relatively low computational and implementation complexity when compared to other

**Original** **Noisy**

**NS [7]** **NLM [22]**

**BM3D [24]** **MCMCD**

Figure 6: "Barbara" image contaminated by additive Gaussian noise with $\sigma_n = 30$ and the corresponding denoised images using the tested methods.

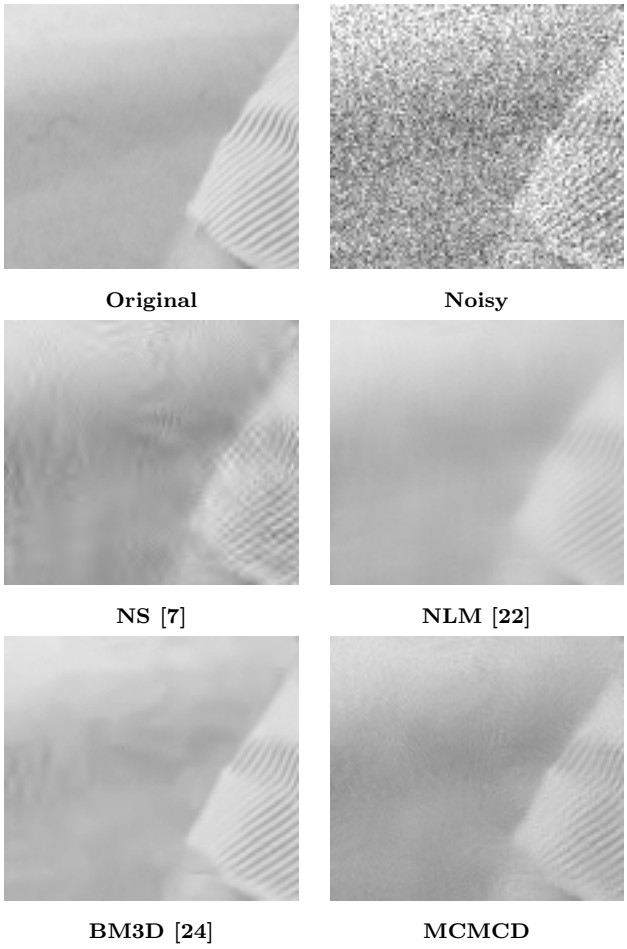



Figure 7: Zoomed-in regions of "Barbara" image contaminated by additive Gaussian noise with $\sigma_n = 30$ and the corresponding denoised images using the tested methods. NS produced results with significant artifacts that degrade the overall visual quality of the image. NLM produced results with significantly fewer artifacts than NS, but does not preserve fine details such as the stripe patterns in Barbara's clothes as well as BM3D and MCMCD. BM3D and MCMCD produced visually comparable results with significant fewer artifacts than NS as well as with better image detail preservation than NLM. Finally, BM3D provides slightly better detail preservation while MCMCD exhibits slightly weaker artifacts in the homogeneous regions.

**Original** **Noisy**

**NS [7]** **NLM [22]**

**BM3D [24]** **MCMCD**

Figure 8: "Lena" image contaminated by additive Gaussian noise with $\sigma_n = 40$ and the corresponding denoised images using the tested methods.

published methods with comparable denoising performance. Experimental results demonstrate the effectiveness of MCMCD in achieving strong image denoising performance when compared to state-of-the-art methods. Future work involves extending the algorithm for multichannel images, video, and 3D data volumes.

## Acknowledgment

This research has been sponsored by the Natural Sciences and Engineering Research Council (NSERC) of Canada through individual Discovery Grants as well as GEOIDE (GEOmatics for Informed Decisions) which is a Network of Centres of Excellence under NSERC. The authors would also like to thank Zhou Dengwen and Kostadin Dabov et al. for making available their source code for testing.

## References

[1] N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, New York: Wiley, 1949.

[2] L. Yaroslavsky, Local adaptive image restoration and enhancement with the use of DFT and DCT in a running window, Proceedings of Wavelet applications in signal and image processing IV, SPIE Proc. Series, 2825 (1996) 1-13.

[3] J. Portilla, V. Strela, M. Wainwright, M., and E. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Transactions on Image Processing, 12(11) (2003) 1338-1351.

[4] Q. Li and C. He, Application of wavelet threshold to image de-noising, Proceedings of Innovative Computing, Information and Control, 2 (2006) 693-696.

[5] H. Biao, J. Li-cheng, L. Fang, Image denoising based on ridgelet, Proceedings of 6th International Conference on Signal Processing, 1 (2002) 780-783.

[6] F. Luisier, T. Blu, and M. Unser, A new SURE approach to image denoising: interscale orthonormal wavelet thresholding, IEEE Transactions on Image Processing, 16(3) (2007) 593-606.

[7] D. Zhou and W. Cheng, Image denoising with an optimal threshold and neigbhouring window, Pattern Recognition Letters, 29 (2008) 1694-1697.

[8] M. Malfait and D. Roose, Wavelet-Based Image Denoising Using a Markov Random Field a Priori Model, IEEE Transactions on Image Processing, 6 (1997) 549-565.

[9] A. Pizurica, W. Philips, I. Lemahieu, and M Acheroy, A Joint Inter- and Intrascale Statistical Model for Bayesian Wavelet Based Image Denoising, IEEE Transactions on Image Processing, 11 (2002) 545-557.

[10] A. Foi, V. Katkovnik, and K. Egiazarian, Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color Images, IEEE Transactions on Image Processing, 16(5) (2007), 1395-1411.

[11] M. Lindenbaum, M. Fischer, and A. Bruckstein, On gabor contribution to image- enhancement, Pattern Recognition, 27 (1994) 1-8.

[12] J. Lee, Digital image enhancement and noise filtering by use of local statistics, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2 (1980) 165-168.

[13] L. Rudin and S. Osher, Total variation based image restoration with free local constraints, Proceedings of IEEE International Conference on Image Processing, 1 (1994) 31-35.

[14] S. Geman and D. Geman, Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Trans. Pattern Anal. Machine Intel., PAMI-6 (1984) 721-741.

[15] J. Pryce and A. Bruce, Statistical mechanics of image restoration, J. Phys. A: Math. Gen., 28(3) (1995) 511-532.

[16] S. Greenberg and D. Kogan, Improved structure-adaptive anisotropic filter, Pattern Recognition Letters, 27(1) (2006), 59-65.

[17] C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, Proceedings of IEEE International Conference on Computer Vision, (1998) 836-846.

[18] M. Elad, On the origin of the bilateral filter and ways to improve it, IEEE Transactions on Image Processing, 11(10) (2002) 1141-1151.

[19] A. Wong, Adaptive bilateral filtering of image signals using local phase characteristics, Signal Processing, 88(6) (2008) 1615-1619.

[20] W. Wong, A. Chung, and S. Yu, Trilateral filtering for biomedical im-

ages, Proceedings of IEEE International Symposium on Biomedical Imaging: Nano to Macro, 1 (2004) 820-823.

[21] M. Mahmoudi and G. Sapiro, Fast image and video denoising via nonlocal means of similar neighborhoods, IEEE Signal Processing Letters, 12(12) (2005) 839-842.

[22] A. Buades, B. Coll, and J. Morel, Nonlocal image and movie denoising, International Journal of Computer Vision, 76(2) (2008) 123-139.

[23] J. Orchard, M. Ebrahimi, and A. Wong, Efficient nonlocal-means denoising using the SVD, Proceedings of IEEE International Conference on Image Processing, (2008).

[24] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, Image denoising by sparse 3D transform-domain collaborative filtering, IEEE Transactions on Image Processing, 16(8) (2007) 2080-2095.

[25] P. Fieguth, Statistical Image Processing and Multidimensional Modeling, first ed., Springer, New York, 2010.

[26] J. S. Lee, Speckle suppression and analysis for synthetic aperture radar, Opt. Eng., 25(5) (1986) 636-643.

[27] A. E. Lopes, E. Nezry, R. Touzi, and H. Laur, Structure Detection and Adaptive Speckle Filtering in SAR Images, Int. J. Remote Sens., 14(9) (1993) 1735-1758.

[28] W. Hastings, Monte carlo sampling methods using Markov chains and their applications, Biometrika, 57(1) (1970) 97-109.

[29] S. Geman and D. McClure, Statistical methods for tomographic image reconstruction, Bulletin of the International Statistical Institute, LII-4 (1987) 5-21.

[30] M. Chen, Importance-weighted marginal Bayesian posterior density estimation, Journal of the American Statistical Associtation, 89(427) (1994) 818-824.

[31] The USC-SIPI Image Database, http://sipi.usc.edu/database.

[32] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, 13(4) (2004) 600-612.